

Снижение размерности пространства в задачах анализа временных рядов

Исаченко Роман Владимирович

Диссертация на соискание ученой степени
кандидата физико-математических наук

05.13.17 – Теоретические основы информатики

Научный руководитель: д.ф.-м.н. В. В. Стрижов

Москва, 2021 г.

Снижение размерности пространства в задачах анализа временных рядов

Цель

Исследовать зависимости в пространствах объектов и ответов и построить устойчивую модель декодирования временных рядов в случае коррелированного описания данных.

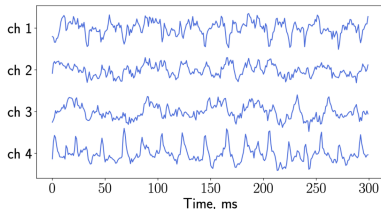
Проблема

Целевая переменная – вектор, компоненты которого являются зависимыми. Требуется построить модель, адекватно описывающую как пространство объектов так и пространство ответов при наблюдаемой мультикорреляции в обоих пространствах высокой размерности.

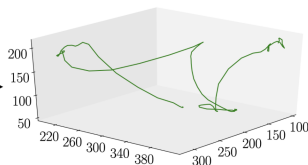
Решение

Для учёта зависимостей в пространствах объектов и ответов предлагается снизить размерность с использованием скрытого пространства. Предлагаются линейные и нейросетевые методы согласования связанных моделей в пространствах высокой размерности.

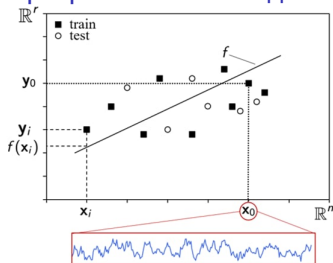
Задача декодирования сигналов



f



Авторегрессионная модель

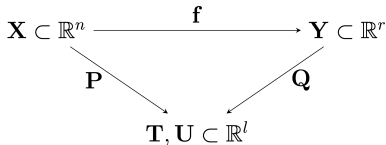


Проекция в скрытое пространство

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{F}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{E}$$

$$\text{cov}(\mathbf{T}, \mathbf{U}) \rightarrow \max_{\mathbf{P}, \mathbf{Q}}$$



Задача авторегрессионного прогнозирования

$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ – матрица объектов;

$\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_r] \in \mathbb{R}^{m \times r}$ – матрица ответов;

$\mathbf{y} = \mathbf{\Theta} \mathbf{x} + \epsilon$, $\mathbf{\Theta} \in \mathbb{R}^{r \times n}$ – модель.

Функция потерь линейной регрессии

$$\mathcal{L}(\mathbf{\Theta} | \mathbf{X}, \mathbf{Y}) = \left\| \begin{matrix} \mathbf{Y} & \mathbf{X} & \mathbf{\Theta}^T \\ m \times r & m \times n & r \times n \end{matrix} \right\|_2^2 \rightarrow \min_{\mathbf{\Theta}} \Rightarrow \mathbf{\Theta}^T = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}.$$

Для устранения сильной линейной зависимости столбцов матрицы \mathbf{X} предлагается использовать методы снижения размерности пространства.

Метод проекции в скрытое пространство (PLS)

$$\mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{F} = \sum_{k=1}^l \mathbf{t}_k \cdot \mathbf{p}_k^T + \mathbf{F},$$

$m \times n \quad m \times l \quad l \times n \quad m \times n \quad m \times 1 \quad 1 \times n \quad m \times n$

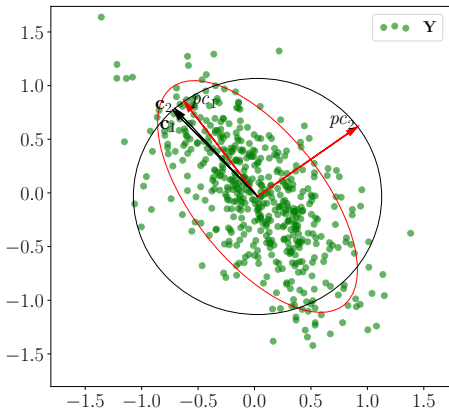
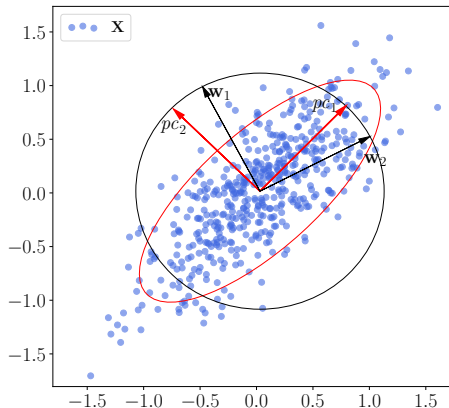
$$\mathbf{Y} = \mathbf{U} \cdot \mathbf{Q}^T + \mathbf{E} = \sum_{k=1}^l \mathbf{u}_k \cdot \mathbf{q}_k^T + \mathbf{E}.$$

$m \times r \quad m \times l \quad l \times r \quad m \times r \quad m \times 1 \quad 1 \times r \quad m \times r$

$$\hat{\mathbf{Y}} = f(\mathbf{X}) = \mathbf{X} \mathbf{\Theta}; \quad \mathbf{U} \approx \mathbf{T} \mathbf{B}, \quad \mathbf{B} = \text{diag}(\beta_k), \quad \beta_k = \mathbf{u}_k^T \mathbf{t}_k / (\mathbf{t}_k^T \mathbf{t}_k).$$

Пример PLS регрессии в двумерном случае

- $\mathbf{x}_i \sim \mathcal{N}(0, \Sigma)$;
- \mathbf{y}_i линейно зависят от pc_2 и не зависят от pc_1 .



Учёт взаимной связи между матрицами \mathbf{X} и \mathbf{Y} отклоняет вектора \mathbf{w}_k и \mathbf{c}_k от направления главных компонент.

Метод частных наименьших квадратов (PLS)

Утверждение (Исаченко, 2017)

Максимизация ковариации между векторами \mathbf{t}_k и \mathbf{u}_k приводит к наилучшему описанию матриц \mathbf{X} и \mathbf{Y} с учётом их взаимосвязи.

Утверждение (Исаченко, 2017)

Вектора \mathbf{w}_k и \mathbf{c}_k – собственные вектора матриц $\mathbf{X}_k^T \mathbf{Y}_k \mathbf{Y}_k^T \mathbf{X}_k$ и $\mathbf{Y}_k^T \mathbf{X}_k \mathbf{X}_k^T \mathbf{Y}_k$, соответствующие максимальным собственным значениям.

Утверждение (Исаченко, 2017)

Правила обновления векторов (6)–(9) максимизируют ковариацию между векторами \mathbf{t}_k и \mathbf{u}_k .

Модель PLS регрессии

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{E} \approx \mathbf{TBQ}^T + \mathbf{E} = \mathbf{XW}^* \mathbf{BQ}^T + \mathbf{E} = \mathbf{X\Theta} + \mathbf{E}.$$
$$\mathbf{\Theta} = \mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1} \mathbf{BQ}^T, \quad \mathbf{T} = \mathbf{XW}^*, \quad \text{where } \mathbf{W}^* = \mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1}.$$

Скрытое пространство в задаче декодирования

Особенностью задачи является избыточность описания независимой переменной \mathbf{x} и целевой переменной \mathbf{y} . Объекты \mathbf{x} и \mathbf{y} живут на некоторых многообразиях низкой размерности.

$$\begin{array}{ccc} \mathbf{X} \subset \mathbb{R}^n & \xrightarrow{f} & \mathbf{Y} \subset \mathbb{R}^r \\ \varphi_e \updownarrow \varphi_d & & \psi_d \updownarrow \psi_e \\ \mathbf{T} \subset \mathbb{R}^l & \xrightarrow{h} & \mathbf{U} \subset \mathbb{R}^s \end{array}$$

Пространства $\mathbf{T} \subset \mathbb{R}^l$ и $\mathbf{U} \subset \mathbb{R}^s$ *скрытые пространства* для $\mathbf{X} \in \mathbb{R}^n$ ($l \leq n$) и $\mathbf{Y} \in \mathbb{R}^r$ ($s \leq r$), если существуют *функции кодирования* $\varphi_e : \mathbf{X} \rightarrow \mathbf{T}$, $\psi_e : \mathbf{Y} \rightarrow \mathbf{U}$ и *функции декодирования* $\varphi_d : \mathbf{T} \rightarrow \mathbf{X}$, $\psi_d : \mathbf{U} \rightarrow \mathbf{Y}$:

$$\begin{aligned} \forall \mathbf{x} \in \mathbf{X} \quad \exists \mathbf{t} \in \mathbf{T} : \varphi_d(\varphi_e(\mathbf{x})) &= \varphi_d(\mathbf{t}) = \mathbf{x}; \\ \forall \mathbf{y} \in \mathbf{Y} \quad \exists \mathbf{u} \in \mathbf{U} : \psi_d(\psi_e(\mathbf{y})) &= \psi_d(\mathbf{u}) = \mathbf{y}. \end{aligned}$$

Скрытые пространства \mathbf{T} и \mathbf{U} являются *согласованными*, если существует *функция согласования* $h : \mathbf{T} \rightarrow \mathbf{U}$:

$$\mathbf{y} = f(\mathbf{x}) = \psi_d(h(\varphi_e(\mathbf{x}))).$$

Выбор признаков в задаче декодирования

Требуется

Найти бинарный вектор $\mathbf{a} = \{0, 1\}^n$, компоненты – индикаторы выбранных признаков.

Функция ошибки отбора признаков

$$\mathbf{a} = \arg \min_{\mathbf{a}' \in \{0,1\}^n} S(\mathbf{a}' | \mathbf{X}, \mathbf{Y}).$$

Релаксация

Замена дискретной области определения $\{0, 1\}^n$ на непрерывную релаксацию $[0, 1]^n$:

$$\mathbf{z} = \arg \min_{\mathbf{z}' \in [0,1]^n} S(\mathbf{z}' | \mathbf{X}, \mathbf{Y}), \quad a_j = [z_j > \tau].$$

Получив \mathbf{a} , решаем задачу регрессии:

$$\mathcal{L}(\Theta_{\mathbf{a}} | \mathbf{X}_{\mathbf{a}}, \mathbf{Y}) = \left\| \mathbf{Y} - \mathbf{X}_{\mathbf{a}} \Theta_{\mathbf{a}}^T \right\|_2^2 \rightarrow \min_{\Theta_{\mathbf{a}}},$$

где индекс \mathbf{a} обозначает подматрицу с номерами столбцов, для которых $a_j = 1$.

Выбор признаков с помощью квадратичного программирования

$$\|\nu - \mathbf{X}\theta\|_2^2 \rightarrow \min_{\theta \in \mathbb{R}^n}.$$

Задача квадратичного программирования

$$S(\mathbf{z}|\mathbf{X}, \nu) = (1 - \alpha) \cdot \underbrace{\mathbf{z}^T \mathbf{Q} \mathbf{z}}_{\text{Sim}(\mathbf{X})} - \alpha \cdot \underbrace{\mathbf{b}^T \mathbf{z}}_{\text{Rel}(\mathbf{X}, \nu)} \rightarrow \min_{\substack{\mathbf{z} \geq \mathbf{0}_n \\ \mathbf{1}_n^T \mathbf{z} = 1}}.$$

$\mathbf{z} \in [0, 1]^n$ – значимость признаков;

$\mathbf{Q} \in \mathbb{R}^{n \times n}$ – матрица парных взаимодействий признаков;

$\mathbf{b} \in \mathbb{R}^n$ – вектор релевантностей признаков к целевой переменной.

$$\mathbf{Q} = [\text{corr}(\chi_i, \chi_j)]_{i,j=1}^n, \quad \mathbf{b} = [\text{corr}(\chi_i, \nu)]_{i=1}^n.$$

Утверждение (Исаченко, 2018)

В случае полуопределенной матрицы \mathbf{Q} задача QPFS является выпуклой.
Полуопределенная релаксация – сдвиг спектра:

$$\mathbf{Q} \rightarrow \mathbf{Q} - \lambda_{\min} \mathbf{I}.$$

Многомерный выбор признаков в задаче декодирования

Агрегирование релевантностей по целевым векторам (RelAgg)

$$\mathbf{b} = [|\text{corr}(\chi_i, \nu)|]_{i=1}^n \rightarrow \mathbf{b} = \left[\sum_{k=1}^r |\text{corr}(\chi_i, \nu_k)| \right]_{i=1}^n.$$

Недостаток: нет учёта зависимостей в матрице \mathbf{Y} .

Симметричный учёт значимостей (SymImp)

Штрафуем коррелированные целевые вектора с помощью $\text{Sim}(\mathbf{Y})$

$$\alpha_1 \cdot \underbrace{\mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x}_{\text{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\mathbf{z}_x^T \mathbf{B} \mathbf{z}_y}_{\text{Rel}(\mathbf{X}, \mathbf{Y})} + \alpha_3 \cdot \underbrace{\mathbf{z}_y^T \mathbf{Q}_y \mathbf{z}_y}_{\text{Sim}(\mathbf{Y})} \rightarrow \min_{\substack{\mathbf{z}_x \geq \mathbf{0}_n, \mathbf{1}_n^T \mathbf{z}_x = 1 \\ \mathbf{z}_y \geq \mathbf{0}_r, \mathbf{1}_r^T \mathbf{z}_y = 1}}.$$

$$\mathbf{Q}_x = [|\text{corr}(\chi_i, \chi_j)|]_{i,j=1}^n, \quad \mathbf{Q}_y = [|\text{corr}(\nu_i, \nu_j)|]_{i,j=1}^r, \quad \mathbf{B} = [|\text{corr}(\chi_i, \nu_j)|]_{\substack{i=1,\dots,n \\ j=1,\dots,r}}.$$

$$\alpha_1 + \alpha_2 + \alpha_3 = 1, \quad \alpha_i \geq 0.$$

Многомерный выбор признаков в задаче декодирования

SymImp штрафует коррелированные целевые вектора, которые в меньшей мере объясняются признаками.

$$\alpha_1 \cdot \underbrace{\mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x}_{\text{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\mathbf{z}_x^T \mathbf{B} \mathbf{z}_y}_{\text{Rel}(\mathbf{X}, \mathbf{Y})} \rightarrow \min_{\substack{\mathbf{z}_x \geq \mathbf{0}_n \\ \mathbf{1}_n^T \mathbf{z}_x = 1}} ; \quad \alpha_3 \cdot \underbrace{\mathbf{z}_y^T \mathbf{Q}_y \mathbf{z}_y}_{\text{Sim}(\mathbf{Y})} + \alpha_2 \cdot \underbrace{\mathbf{z}_x^T \mathbf{B} \mathbf{z}_y}_{\text{Rel}(\mathbf{X}, \mathbf{Y})} \rightarrow \min_{\substack{\mathbf{z}_y \geq \mathbf{0}_r \\ \mathbf{1}_r^T \mathbf{z}_y = 1}} .$$

Минимаксный подход (MinMax / MaxMin)

$$\min_{\substack{\mathbf{z}_x \geq \mathbf{0}_n \\ \mathbf{1}_n^T \mathbf{z}_x = 1}} \max_{\substack{\mathbf{z}_y \geq \mathbf{0}_r \\ \mathbf{1}_r^T \mathbf{z}_y = 1}} \left(\text{or } \max_{\substack{\mathbf{z}_y \geq \mathbf{0}_r \\ \mathbf{1}_r^T \mathbf{z}_y = 1}} \min_{\substack{\mathbf{z}_x \geq \mathbf{0}_n \\ \mathbf{1}_n^T \mathbf{z}_x = 1}} \right) \left[\alpha_1 \cdot \underbrace{\mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x}_{\text{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\mathbf{z}_x^T \mathbf{B} \mathbf{z}_y}_{\text{Rel}(\mathbf{X}, \mathbf{Y})} - \alpha_3 \cdot \underbrace{\mathbf{z}_y^T \mathbf{Q}_y \mathbf{z}_y}_{\text{Sim}(\mathbf{Y})} \right] .$$

Теорема (Исаченко, 2018)

Для положительно определенных матриц \mathbf{Q}_x и \mathbf{Q}_y minmax и maxmin задачи достигают одинакового значения функционала.

Теорема (Исаченко, 2018)

Минимаксная задача эквивалентна задаче квадратичного программирования с $n + r + 1$ переменными.

Для получения выпуклой задачи применяется сдвиг спектра.

Многомерный выбор признаков в задаче декодирования

Максимизация релевантностей (MaxRel)

$$\min_{\substack{\mathbf{z}_x \geq \mathbf{0}_n \\ \mathbf{1}_n^\top \mathbf{z}_x = 1}} \max_{\substack{\mathbf{z}_y \geq \mathbf{0}_r \\ \mathbf{1}_r^\top \mathbf{z}_y = 1}} \left[(1 - \alpha) \cdot \mathbf{z}_x^\top \mathbf{Q}_x \mathbf{z}_x - \alpha \cdot \mathbf{z}_x^\top \mathbf{B} \mathbf{z}_y \right].$$

Теорема (Исаченко, 2018)

Для положительно определенной матрицы \mathbf{Q}_x *minmax* и *maxmin* задачи достигают одинакового значения функционала.

Асимметричный учёт значимостей (AsymImp)

$$\alpha_1 \cdot \underbrace{\mathbf{z}_x^\top \mathbf{Q}_x \mathbf{z}_x}_{\text{Sim}(\mathbf{X})} - \alpha_2 \cdot \underbrace{\left(\mathbf{z}_x^\top \mathbf{B} \mathbf{z}_y - \mathbf{b}^\top \mathbf{z}_y \right)}_{\text{Rel}(\mathbf{X}, \mathbf{Y})} + \alpha_3 \cdot \underbrace{\mathbf{z}_y^\top \mathbf{Q}_y \mathbf{z}_y}_{\text{Sim}(\mathbf{Y})} \rightarrow \min_{\substack{\mathbf{z}_x \geq \mathbf{0}_n, \mathbf{1}_n^\top \mathbf{z}_x = 1 \\ \mathbf{z}_y \geq \mathbf{0}_r, \mathbf{1}_r^\top \mathbf{z}_y = 1}}.$$

При $b_j = \max_{i=1, \dots, n} [\mathbf{B}]_{i,j}$ коэффициенты при \mathbf{z}_y в $\text{Rel}(\mathbf{X}, \mathbf{Y})$ неотрицательны.

Утверждение (Исаченко, 2017)

В одномерном случае $r = 1$ предлагаемые стратегии *SymImp*, *MinMax*, *MaxMin*, *MaxRel*, *AsymImp* совпадают с исходным алгоритмом *QPFS*.

Обобщение предложенных методов выбора признаков

Алгоритм	Критерий	Функция ошибки $S(\mathbf{z} \mathbf{X}, \mathbf{Y})$
RelAgg	$\min [\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})]$	$\min_{\mathbf{z}_x} [(1 - \alpha) \cdot \mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x - \alpha \cdot \mathbf{z}_x^T \mathbf{B} \mathbf{1}_r]$
SymImp	$\min [\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y}) + \text{Sim}(\mathbf{Y})]$	$\min_{\mathbf{z}_x, \mathbf{z}_y} [\alpha_1 \cdot \mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x - \alpha_2 \cdot \mathbf{z}_x^T \mathbf{B} \mathbf{z}_y + \alpha_3 \cdot \mathbf{z}_y^T \mathbf{Q}_y \mathbf{z}_y]$
MinMax	$\min [\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})]$ $\max [\text{Rel}(\mathbf{X}, \mathbf{Y}) + \text{Sim}(\mathbf{Y})]$	$\min_{\mathbf{z}_x} \max_{\mathbf{z}_y} [\alpha_1 \cdot \mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x - \alpha_2 \cdot \mathbf{z}_x^T \mathbf{B} \mathbf{z}_y - \alpha_3 \cdot \mathbf{z}_y^T \mathbf{Q}_y \mathbf{z}_y]$
MaxRel	$\min [\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})]$ $\max [\text{Rel}(\mathbf{X}, \mathbf{Y})]$	$\min_{\mathbf{z}_x} \max_{\mathbf{z}_y} [(1 - \alpha) \cdot \mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x - \alpha \cdot \mathbf{z}_x^T \mathbf{B} \mathbf{z}_y]$
AsymImp	$\min [\text{Sim}(\mathbf{X}) - \text{Rel}(\mathbf{X}, \mathbf{Y})]$ $\max [\text{Rel}(\mathbf{X}, \mathbf{Y}) + \text{Sim}(\mathbf{Y})]$	$\min_{\mathbf{z}_x, \mathbf{z}_y} [\alpha_1 \mathbf{z}_x^T \mathbf{Q}_x \mathbf{z}_x - \alpha_2 (\mathbf{z}_x^T \mathbf{B} \mathbf{z}_y - \mathbf{b}^T \mathbf{z}_y) + \alpha_3 \mathbf{z}_y^T \mathbf{Q}_y \mathbf{z}_y]$

Внешние критерии качества

Нормированное RMSE

Качество предсказания:

$$\text{sRMSE}(\mathbf{Y}, \hat{\mathbf{Y}}_a) = \sqrt{\frac{\text{MSE}(\mathbf{Y}, \hat{\mathbf{Y}}_a)}{\text{MSE}(\mathbf{Y}, \bar{\mathbf{Y}})}} = \frac{\|\mathbf{Y} - \hat{\mathbf{Y}}_a\|_2}{\|\mathbf{Y} - \bar{\mathbf{Y}}\|_2}, \quad \text{где} \quad \hat{\mathbf{Y}}_a = \mathbf{X}_a \boldsymbol{\Theta}_a^T.$$

$\bar{\mathbf{Y}}$ — константный прогноз.

Мультикорреляция

Среднее значение коэффициента множественной корреляции:

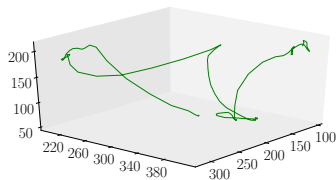
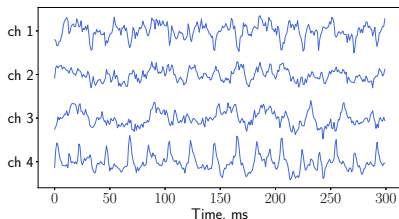
$$R^2 = \frac{1}{r} \text{tr} \left(\mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \right); \quad \mathbf{C} = [\text{corr}(\chi_i, \nu_j)]_{\substack{i=1, \dots, n \\ j=1, \dots, r}}, \quad \mathbf{R} = [\text{corr}(\chi_i, \chi_j)]_{i,j=1}^n.$$

BIC

Компромисс между качеством предсказания и количеством выбранных признаков $\|\mathbf{a}\|_0$:

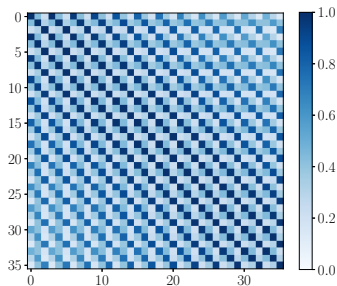
$$\text{BIC} = m \ln \left(\text{MSE}(\mathbf{Y}, \hat{\mathbf{Y}}_a) \right) + \|\mathbf{a}\|_0 \cdot \log m.$$

Данные ECoG



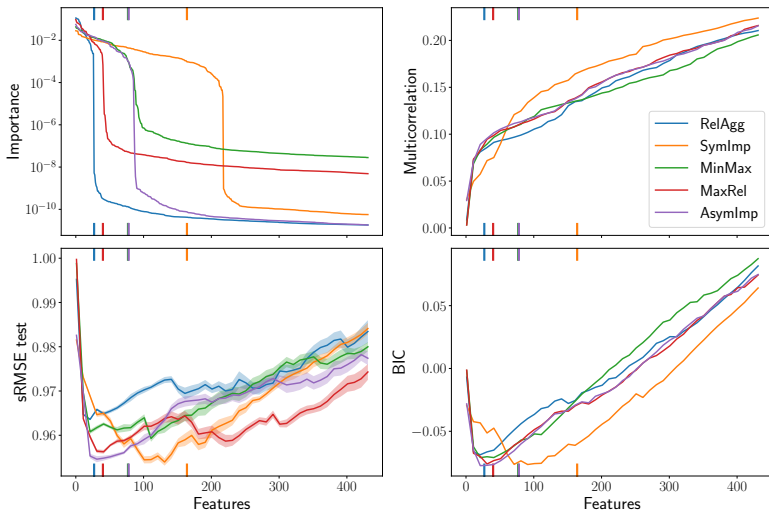
- $\mathbf{X} \in \mathbb{R}^{m \times (32 \cdot 27)}$ – сигналы ECoG.
- $\mathbf{Y} \in \mathbb{R}^{m \times 3k}$ – траектория движения руки.

$$\mathbf{Y} = \begin{pmatrix} x_1 & y_1 & z_1 & \dots & x_k & y_k & z_k \\ x_2 & y_2 & z_2 & \dots & x_{k+1} & y_{k+1} & z_{k+1} \\ \dots & \dots & \dots & & \dots & \dots & \dots \\ x_m & y_m & z_m & \dots & x_{m+k} & y_{m+k} & z_{m+k} \end{pmatrix}$$



Матрица корреляций \mathbf{Y}

Анализ предложенных методов выбора признаков



Предложенные методы выбирают модель с меньшей ошибкой по отношению к базовому алгоритму.

Стабильность методов выбора признаков

Постановка эксперимента

- создать бутстреп-выборки

$$(\mathbf{X}, \mathbf{Y}) \rightarrow \{(\mathbf{X}_1, \mathbf{Y}_1), \dots, (\mathbf{X}_s, \mathbf{Y}_s)\};$$

- решить задачу выбора признаков

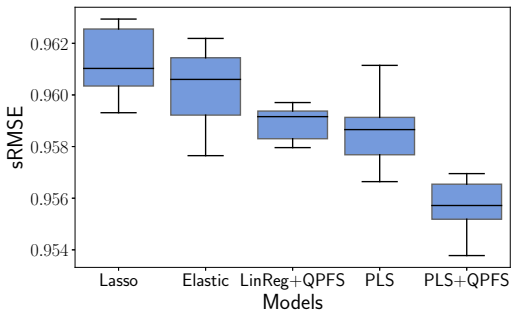
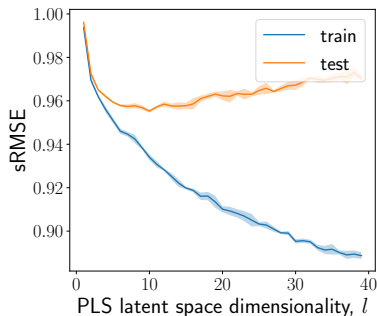
$$\{(\mathbf{X}_1, \mathbf{Y}_1), \dots, (\mathbf{X}_s, \mathbf{Y}_s)\} \rightarrow \{\mathbf{z}_1, \dots, \mathbf{z}_s\};$$

- вычислить статистики

$$\{\mathbf{z}_1, \dots, \mathbf{z}_s\} \rightarrow \{\text{sRMSE}, \|\mathbf{a}\|_0, \text{Спирмен } \rho, \ell_2 \text{ расстояние}\}.$$

	sRMSE	$\ \mathbf{a}\ _0$	Спирмен ρ	ℓ_2 расстояние
RelAgg	0.965 ± 0.002	26.8 ± 3.8	0.915 ± 0.016	0.145 ± 0.018
SymImp	0.961 ± 0.001	224.4 ± 9.0	0.910 ± 0.017	0.025 ± 0.002
MinMax	0.961 ± 0.002	101.0 ± 2.1	0.932 ± 0.009	0.059 ± 0.004
MaxRel	0.958 ± 0.003	41.2 ± 5.2	0.862 ± 0.027	0.178 ± 0.010
AsymImp	0.955 ± 0.001	85.8 ± 10.2	0.926 ± 0.011	0.078 ± 0.007

Сравнение метода проекции в скрытое пространство с методами выбора признаков



- Предлагаемые методы выбора признаков достигают меньшей ошибки по сравнению с базовыми алгоритмами Lasso и Elastic.
- PLS показывает сравнимое качество с QPFS.
- Комбинация двух алгоритмов показывает наилучший результат.

Нелинейные методы согласования скрытого пространства

Нелинейная проекция в скрытое пространство

$$\mathbf{T} = \varphi_e(\mathbf{X}) = \mathbf{W}_x^L \sigma(\dots \sigma(\mathbf{W}_x^2 \sigma(\mathbf{X} \mathbf{W}_x^1)) \dots)$$

$$\mathbf{U} = \psi_e(\mathbf{Y}) = \mathbf{W}_y^L \sigma(\dots \sigma(\mathbf{W}_y^2 \sigma(\mathbf{Y} \mathbf{W}_y^1)) \dots)$$

$$\mathbf{X} = \varphi_d(\mathbf{T}) = \mathbf{W}_t^L \sigma(\dots \sigma(\mathbf{W}_t^2 \sigma(\mathbf{T} \mathbf{W}_t^1)) \dots)$$

$$\mathbf{Y} = \psi_d(\mathbf{U}) = \mathbf{W}_u^L \sigma(\dots \sigma(\mathbf{W}_u^2 \sigma(\mathbf{U} \mathbf{W}_u^1)) \dots)$$

$$\begin{array}{ccc} \mathbf{X} \subset \mathbb{R}^n & \xrightarrow{f} & \mathbf{Y} \subset \mathbb{R}^r \\ \varphi_e \updownarrow \varphi_d & & \psi_d \updownarrow \psi_e \\ \mathbf{T} \subset \mathbb{R}^l & \xrightarrow{h} & \mathbf{U} \subset \mathbb{R}^s \end{array}$$

Данные рукописных цифр



Согласование проекций

$$g(\mathbf{T}, \mathbf{U}) \rightarrow \max_{\mathbf{W}}$$

$$\text{где } \mathbf{W} = \{\{\mathbf{W}_x^i\}_{i=1}^L, \{\mathbf{W}_y^i\}_{i=1}^L, \{\mathbf{W}_t^i\}_{i=1}^L, \{\mathbf{W}_u^i\}_{i=1}^L\}.$$

Результаты, выносимые на защиту

1. Исследована задача декодирования сигналов в пространствах высокой размерности. Исследованы методы снижения размерности с анализом структуры пространства.
2. Предложены методы для выбора признаков, учитывающие зависимости как в пространстве объектов, так и в пространстве ответов.
3. Предложена комбинация методов выбора признаков и снижения размерности пространства. Предложенные алгоритмы выбора признаков доставляют устойчивые и адекватные решения в коррелированных пространствах высокой размерности.
4. Предложены нелинейные методы согласования скрытых пространств для данных со сложно организованной целевой переменной.
5. Создан макет системы, пригнозирующей сигналы в пространстве большой размерности.

Заключение

Публикации ВАК

1. Исаченко Р.В., Стрижов В. В. Метрическое обучение в задачах мультиклассовой классификации временных рядов *Информатика и её применения*, 10(2), 2016.
2. Isachenko R. et al. Feature Generation for Physical Activity Classification. *Artificial Intelligence and Decision Making*, 3, 2018.
3. Isachenko R., Strijov V. Quadratic programming optimization for Newton method. *Lobachevskii Journal of Mathematics*, 39(9), 2018.
4. Isachenko R., Vladimirova M., Strijov V. Dimensionality Reduction for Time Series Decoding and Forecasting Problems. *DEStech Transactions on Computer Science and Engineering*, optim, 2018.
5. Исаченко Р.В., Яушев Ф.Р., Стрижов В.В. Модели согласования скрытого пространства в задаче прогнозирования // Системы и средства информатики, 31(1), 2021.

Выступления с докладом

1. Ломоносов, 2016. Метрическое обучение в задачах мультиклассовой классификации временных рядов.
2. Intelligent Data Processing Conference, 2016. Multimodel forecasting multiscale time series in internet of things.
3. Математические методы распознавания образов, 2017. Локальные модели для классификации объектов сложной структуры.
4. Intelligent Data Processing Conference, 2018, Dimensionality reduction for multicorrelated signal decoding with projections to latent space.
5. Intelligent Data Processing Conference, 2020, Снижение размерности в задаче декодирования временных рядов.